REMARKS

This communication is in response to the Office Action mailed on August 12, 2004. In the Office Action, claims 1-43 were pending.

Claims 22 and 43 were objected to because of lacking antecedent basis for "building an N-gram language model". With this amendment, claim 21 has been amended to depend from claim 20 and claim 42 has been amended to depend from claim 41. These amendments provide clear antecedent basis for the features recited in claims 22 and 43. Accordingly, withdrawal of this objection is respectfully requested.

Claims 1-43 were rejected under 35 U.S.C. 103(a) as being unpatentable over Meteer et al. (ICASSP 93) in view of Masataki (0-8186-7919-0/97IEEE 1997). Many of the claims of the present invention recite features for building a language model using a task independent corpus. A task independent corpus is used since they are readily available and thus saves time and development costs associated with a task dependent corpus. A task independent corpus is typically used to generate general language processing applications. However, the present invention utilizes the task dependent corpus for a specific application. Applicants submit that the features of the present invention recited in the claims are neither taught nor suggested by the prior art.

Meteer et al. describe statistical language modeling that combines N-gram models and context free grammars. Task dependent training text and a phrase-structure grammar are used to develop a recognition grammar. Masataki describes using a task independent N-gram and target task data to develop an N-gram. The task independent N-gram is used as a-priori knowledge and data of the target task is used as a-posteriori knowledge.

The Office Action maintains that combining Masataki's task independent corpus and Meteer's N-gram and context free grammar modeling is prima facie obvious to one of ordinary skill in the art. The Office Action explains that the combination would provide a unified language model, a more accurate model and a

model that increases functionality. However, these assertions are not provided with any evidence and are simply inaccurate. As discussed above, task dependent corpora are costly to develop. Meteer has chosen to use a task development corpus to generate a unified language model that is accurate and functional, not to increase cost. Meteer incurred the excess costs inherent with a task dependent corpus to develop a language model. It simply does not make sense that a task independent corpus could be substituted into the system of Meteer. The Office Action does not provide substantive evidence that a more accurate and functional language model would result from this combination. In fact, using the techniques of Meteer with a task independent corpus would likely reduce reliability and accuracy of the resulting language model. As a result, applicants maintain that there is no teaching or suggestion to combine the modeling of Meteer and the task independent corpus of Masataki.

Even if the references are combined, for sake of argument, the features of the claims of the present invention are not taught or suggested. Independent claim 1 recites a method for creating a task dependent unified language model for a selected application from a task independent corpus. The task dependent unified language model has embedded context-free grammar non-terminal tokens in a N-gram model. The method includes obtaining a plurality of context-free grammars comprising non-terminal tokens representing semantic or syntactic concepts. Each of the context-free grammars has words present in the task independent corpus to form the semantic or syntactic concepts. The task independent corpus is parsed with a plurality of context-free grammars to identify word occurrences of each of the semantic or syntactic concepts. Each of the identified word occurrences is replaced with corresponding non-terminal tokens and an N-gram model is built having the non-terminal tokens. The method further includes obtaining a second plurality of context-free grammars comprising at least some of the same non-terminals representing the same semantic or syntactic concepts. Each of the context-free

grammars of the second plurality is more appropriate for use in the selected application. Independent claim 23 recites a computer readable medium having similar features to that of claim 1.

On page 3, the Office Action asserts that Meteer et al. describe a plurality of context-free grammars, citing page II-38, left column, paragraph 3. On page 4, the Office Action further asserts that Meteer et al. teach a second plurality of context free grammars, also referring to page II-38, left column, paragraph 3. Applicants are not clear as to how the same set of rules teaches both a first plurality of context-free grammars and a second plurality of context-free grammars. The Office Action asserts that Meteer et al. use a subset of rules, but the subset referred to is merely to show an example of the phrase grammar of FIG. 1. In fact, FIG. 1 clearly shows that the phase grammar is used both with the sparse parsing and with the construction of the recognition grammar. Thus, it appears that the Office Action has not asserted prima facie evidence of using separate sets of context-free grammars, one of which is more appropriate for a selected application. Instead, Meteer et al. disclose using the same phrase-grammar throughout the language model construction process. Page 5 of the Office Action asserts that a combination of Masataki's task independent corpus and Meteer's et al. corpus parsing lead to a unified language model; however, neither Meteer et al. or Masataki teach or suggest the use of a second plurality of context-free grammars.

Independent claim 3 recites another method for creating a task dependent unified language model for a selected application from a task independent corpus. The method includes obtaining at least one context-free grammar having a non-terminal token for a phrase that can be mistaken for one of the desired task dependent semantic or syntactic concepts. Independent claim 24 recites a computer readable medium having features similar to that of claim 3.

On page 6, the Office Action asserts that Meteer et al. uses a grammar from phrases that are replaced with non-terminals

in which the grammars have a semantic structure. This explanation simply does not teach or suggest having a context-free grammar that includes a non-terminal token for a phrase that can be mistaken for one of the desired task dependent semantic or syntactic concepts. On page 21 of Applicants' Specification, it is explained that a context-free grammar is inserted for a person having the last name "Friday", wherein this could easily be mistaken for a day of the week. Neither Meteer et al. nor Masataki teach or suggest this feature. Furthermore, no tangible evidence has been presented in the Office Action to teach this feature.

Independent claim 7 recites a method for creating language model for a selected application from a task independent corpus. The language model is for use in a language processing system. The method includes obtaining a plurality of context-free grammar comprising non-terminal tokens representing semantic or syntactic concepts of the selected application. Furthermore, word phrases are generated from the plurality of context-free grammars and an information retrieval query is formulated from at least one of the word phrases. The task independent corpus is queried based on the query formulated and associated text in the task independent corpus is identified based on the query. A language is then built with the identified text. Independent claim 28 recites a computer readable medium having features similar to that in claim 7.

Additionally, independent claim 14 also recites a method for creating a language model for a selected application from a task independent corpus. In claim 14, a second N-gram language model is built from the identified text. Furthermore, the first N-gram language model and the second N-gram language model are combined to form a third N-gram language model. Independent claim 35 recites a computer readable medium having features similar to that recited in claim 14.

It is respectfully submitted that the combination of Meteer and Masataki simply fail to teach or suggest querying a

task independent corpus based on a formulated query and identifying associated text in the task independent corpus based on the query. Thus, these features recited in independent claim 7, 14, 28 and 35 are simply not taught nor suggested. Furthermore, combining first and second N-gram language models as recited in claims 14 and 35 are neither taught nor suggested.

Independent claim 19 recites a method for creating a unified language model for a selected application from a corpus. The method includes obtaining a plurality of context-free grammars comprising non-terminal tokens representing semantic or syntactic concepts of the selection application. A word language model is built from the corpus and probabilities are assigned to words of at least some of the context-free grammars as of a function of corresponding probabilities obtained for the same words from the word language model. The probabilities are normalized from the language model in each of the context-free grammars as a function of the words allowed by the corresponding context-free grammar. Claim 40 recites a computer readable medium having similar features to those recited in claims 19.

The Office Action asserts that Masataki describes using a coefficient that normalizes a function. However, Masataki does not teach or suggest using a context-free grammar, and thus it is unclear how the coefficient of Masataki is combined with the context-free grammar of Meteer. There simply is no teaching or suggestion to normalize probability values of words in a context-free grammar as a function of words allowed by a corresponding context-free grammar.

In view of the foregoing, applicants respectfully submit that claims 1-43 are in condition for allowance. Favorable action and allowance of all claims is requested.

The Director is authorized to charge any fee deficiency required by this paper or credit any overpayment to Deposit Account No. 23-1123.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: _____

Todd R. Fronek, Reg. No. 48,516
Suite 1600 - International Centre
900 Second Avenue South
Minneapolis, Minnesota 55402-3319
Phone: (612) 334-3222 Fax: (612) 334-3312